NUI Galway
OÉ Gaillimh

Semester 1 Examinations 2014 / 2015

Exam Code(s)	4BCT1, 1EM1, 1SWB1, 4CS1
	4th Year Computer Science and Information Technology
	4th Year Computer Studies
	Erasmus
	Science Without Borders
Module Code(s)	CT422
Module(s)	Modern Information Management
Paper No.	
External Examiner(s)	Dr J. Cox
Internal Examiner(s)	Professor G. Lyons
	Dr M. Madden
	Dr Colm O'Riordan*

Instructions: Answer **any 3** questions; all questions carry equal marks

Duration	2 hours
No. of Pages	3 including cover page
Discipline(s)	Information Technology
Requirements	None

PTO

Q.1.

(a) What is meant by *collaborative filtering*? Illustrate, with a suitable example, how a collaborative filtering approach could be used to make a prediction for a given user for a given item. (10)

(b) Identify the main limitations of a collaborative filtering system and suggest mechanisms to overcome these limitations. (10)

(c) Discuss metrics to evaluate the quality of a collaborative filtering system and discuss the suitability of these metrics. (5)

Q.2.

(a) The performance of an Information Retrieval system is dependent on the quality of the term-weighting scheme employed. With reference to any well-known weighting scheme, discuss the main factors considered in assigning a weight to a term in a document. (9)

(b) Explicit user feedback regarding the quality of the returned answer set is often used to modify the user's query to improve the performance of the retrieval system. Assuming a vector space representation, show how you could modify a user's query when the user can give feedback on returned documents with a value in the range 1 to 5, 1 indicating not relevant at all and 5 indicating very relevant. (7)

(c) Users' queries can be extended by identifying related terms and adding them to the query. Suggest mechanisms to identify related terms. (9)

Q.3.

(a) Discuss applications in information retrieval where clustering algorithms can be of benefit. (9)

(b) Given a set of documents (represented as term-dimensional vectors), outline any algorithm of your choice to cluster the documents together. Discuss the limitations of your approach. (7)

(c) Discuss, briefly, methods that allow the prediction of query difficulty, i.e., any approaches that allow us to identify queries that may be difficult to satisfy correctly. (9)

Q.4.

(a) Describe, with an appropriate example, how the Page Rank algorithm exploits the link structure of a graph to assign a rank to each individual node. Discuss any limitations of the approach. (10)

(b) Suggest modifications to the Page Rank algorithm that may improve the performance of the approach in a web search environment for individual users with specific ongoing interests. (5)

(c) Many different user activities may be involved in information retrieval systems; for example, these activities may include browsing, querying, viewing results etc. Discuss, with reference to existing approaches the issues involved in designing an interface to aid users with these types of activities. (10)